

## Chapter VI

# Content-Based Image Retrieval: From the Object Detection/Recognition Point of View

**Ming Zhang**

*University of Calgary, Canada*

**Reda Alhaji**

*University of Calgary, Canada*

### ABSTRACT

*Content-Based Image Retrieval (CBIR) aims to search images that are perceptually similar to the query-based on visual content of the images without the help of annotations. The current CBIR systems use global features (e.g., color, texture, and shape) as image descriptors, or use features extracted from segmented regions (called region-based descriptors). In the former case, descriptors are not discriminative enough at the object level and are sensitive to object occlusion or background clutter, thus fail to give satisfactory result. In the latter case, the features are sensitive to the image segmentation, which is a difficult task in its own right. In addition, the region-based descriptors are still not invariant to varying imaging conditions. In this chapter, we look at the CBIR from the object detection/recognition point of view and introduce the local feature-based image representation methods recently developed in object detection/recognition area. These local descriptors are highly distinctive and robust to imaging condition change. In addition to image representation, we also introduce the other two key issues of CBIR: similarity measurement for image descriptor comparison and the index structure for similarity search.*

## 1. INTRODUCTION

The explosive growth of digital images in our lives requires efficient image data management systems for image storage and retrieval. The early image retrieval systems are based on manually annotated descriptions and have the following drawbacks (Chang and Hsu, 1992): 1. manually annotating is too expensive for large database; 2. the annotation is subjective and context-dependent. From the early 1990s, **content-based image retrieval (CBIR)** became an active and fast developing research area. Simply speaking, **CBIR** is a method of image retrieval that searches images (from the database) that are similar to the query image based on visual content (by “appearance” according to human perception). More formally, we may define the CBIR as follows:

**Definition of content-based image retrieval:** Given a large image database  $U$ , an image representation method based on image primitives (e.g., pixel intensities) and a dissimilarity measure  $D(p,q)$  defined on the image representation, find (using certain index) the  $M$  images  $p \in U$  with the lowest dissimilarity to the query image  $q$ , the resulting  $M$  images are ranked by ascending dissimilarity.

According to the above definition, the **CBIR** is query by example. Our definition is narrower than the general case **CBIR** as query by example is the most common form of **CBIR**. Our chapter is based on this definition.

As the ultimate goal of image retrieval system is to find images that the users are interested in and the result images are determined by the content of the query image, the first problem we have to deal with is: Can the query image always express clearly what the users are interested in?

In a traditional database, a query is a formally phrased information request that clearly expresses the user’s information needs. Put it in the context of CBIR, the user’s information request should be unambiguously determined by the visual content of the query image. If a user is looking for a specific object or the objects that are very similar to a specific object by appearance, for example, his lost dog or the dogs that look very much like his dog, he may use a photo of his dog as the query image. In this case, the photo is better than any words to express clearly the user’s interest. However, users who are only interested in a generic category of objects, for instance, “animal”, can never be able to express this information need by submitting an image of a specific dog. They will not get satisfactory result by doing so. Some researchers refer to this as the semantic gap problem. To bridge the semantic gap is to make the system generalize semantic concept from a single specific image, which is, in our opinion, impossible, because a high level concept can only be generalized from a large number of instances using machine learning techniques.

In this chapter, we will not consider the semantic gap problem. We assume that the CBIR is used when the users’ information need is better expressed by image than by words. Specifically, query by example is preferred when the user is looking for a specific object/scene or a narrow category of objects/scenes that the instances belonging to this category are similar by appearance, in both cases the user’s information need cannot be expressed unambiguously by simple words. From this perspective, the content-based image retrieval can be roughly regarded as an object detection/recognition problem. Therefore, we may use some techniques originally developed for object detection/recognition in CBIR system.

In this chapter, we discuss three key issues of the CBIR:

28 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/content-based-image-retrieval/4154](http://www.igi-global.com/chapter/content-based-image-retrieval/4154)

## Related Content

---

### The Role of Significant Digits of Fractal Dimension for the Measurement of Texture in SAR Images

Triloki Pant (2021). *International Journal of Computer Vision and Image Processing* (pp. 80-96).

[www.irma-international.org/article/the-role-of-significant-digits-of-fractal-dimension-for-the-measurement-of-texture-in-sar-images/270877](http://www.irma-international.org/article/the-role-of-significant-digits-of-fractal-dimension-for-the-measurement-of-texture-in-sar-images/270877)

### Machine Vision Based Non-Magnetic Object Detection and Removal on Moving Conveyors in Steel Industry through Differential Techniques

K. C. Manjunatha, H. S. Mohanaand P. A. Vijaya (2012). *International Journal of Computer Vision and Image Processing* (pp. 59-70).

[www.irma-international.org/article/machine-vision-based-non-magnetic/74801](http://www.irma-international.org/article/machine-vision-based-non-magnetic/74801)

### Object-Oriented Logic Programming of Intelligent Visual Surveillance for Human Anomalous Behavior Detection

Alexei Alexandrovich Morozov, Olga Sergeevna Sushkovaand Alexander Fedorovich Polupanov (2019). *Optoelectronics in Machine Vision-Based Theories and Applications* (pp. 134-187).

[www.irma-international.org/chapter/object-oriented-logic-programming-of-intelligent-visual-surveillance-for-human-anomalous-behavior-detection/209832](http://www.irma-international.org/chapter/object-oriented-logic-programming-of-intelligent-visual-surveillance-for-human-anomalous-behavior-detection/209832)

### Moving Vehicles Detection in Traffic Video Using Modified SXCS-LBP Texture Descriptor

Arun Kumar H. D.and Prabhakar C. J. (2015). *International Journal of Computer Vision and Image Processing* (pp. 14-34).

[www.irma-international.org/article/moving-vehicles-detection-in-traffic-video-using-modified-sxcs-lbp-texture-descriptor/159805](http://www.irma-international.org/article/moving-vehicles-detection-in-traffic-video-using-modified-sxcs-lbp-texture-descriptor/159805)

### Securing Digital Image from Malicious Insider Attacks

Rahul Saranjameand Manik Lal Das (2018). *International Journal of Computer Vision and Image Processing* (pp. 49-58).

[www.irma-international.org/article/securing-digital-image-from-malicious-insider-attacks/208177](http://www.irma-international.org/article/securing-digital-image-from-malicious-insider-attacks/208177)